



Ecosystems Research Division

## 3MRA Uncertainty and Sensitivity Analysis

SAB Review 3MRA Version 1.0

Panel Meeting: August 26, 2003

Justin Babendreier Office of Research and Development, U.S. EPA

#### **Presentation Outline**

- Modeling System Overview; 3MRA Versions
- 3MRA Version 1.0
- National-scale Assessment Dimensionality
- SuperMUSE: Windows-based Supercomputer
- 3MRA Ver1.x UA/SA Software Tools
- Model Evaluation Approaches
- 3MRA Version 1.0 UA/SA Plan
- Example Model Output Using Version 1.x





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

# Modeling System Version Overview:

FRAMES and 3MRA

# Definition and Relationship between FRAMES and 3MRA

**FRAMES** 

Framework for Risk Analysis in Multimedia Environmental Systems

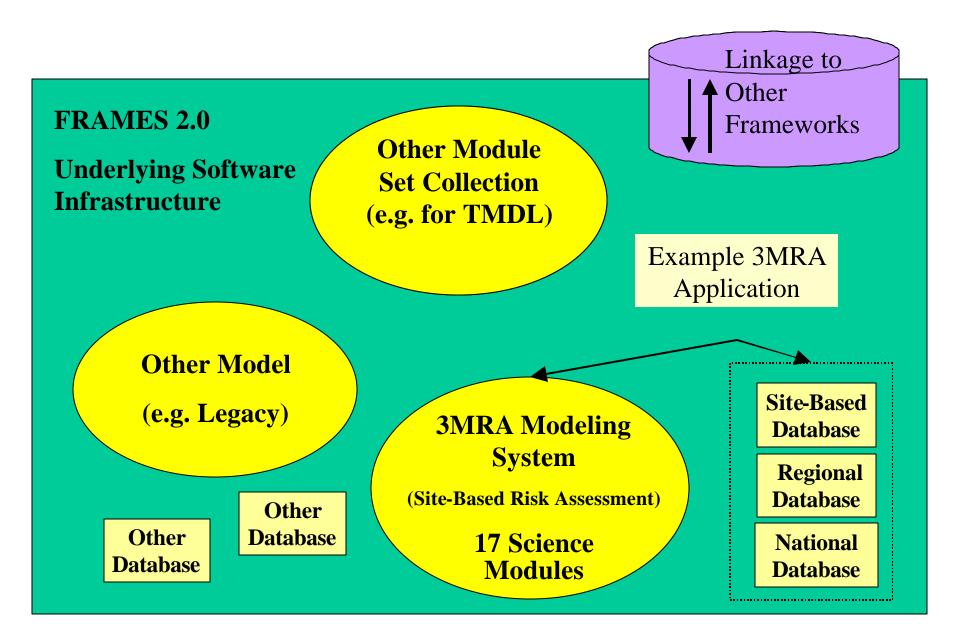
Underlying software infrastructure for 3MRA and other models and modeling systems

**3MRA** 

Multimedia, Multipathway, Multireceptor, Risk Assessment

A specific set of models for conducting site-specific or site-based risk assessments, and "rolled-up" studies on regional and national scales.

#### Conceptual Relationship Between Framework Technology, Models and Modeling System, and Applications



#### FRAMES 3MRA Versions

**3MRA Version 1.0:** National site-based risk assessments

**3MRA Version 1.x:** A tool set extension to facilitate:

- (1) Parallel processing of 3MRA model runs
- (2) Uncertainty and sensitivity analyses studies

3MRA Version 2.0 Beta: Same science and data with an extension to facilitate site-specific risk assessments. Significantly advances the design of the underlying FRAMES infrastructure.

FRAMES 2.0 joint, multi-agency development





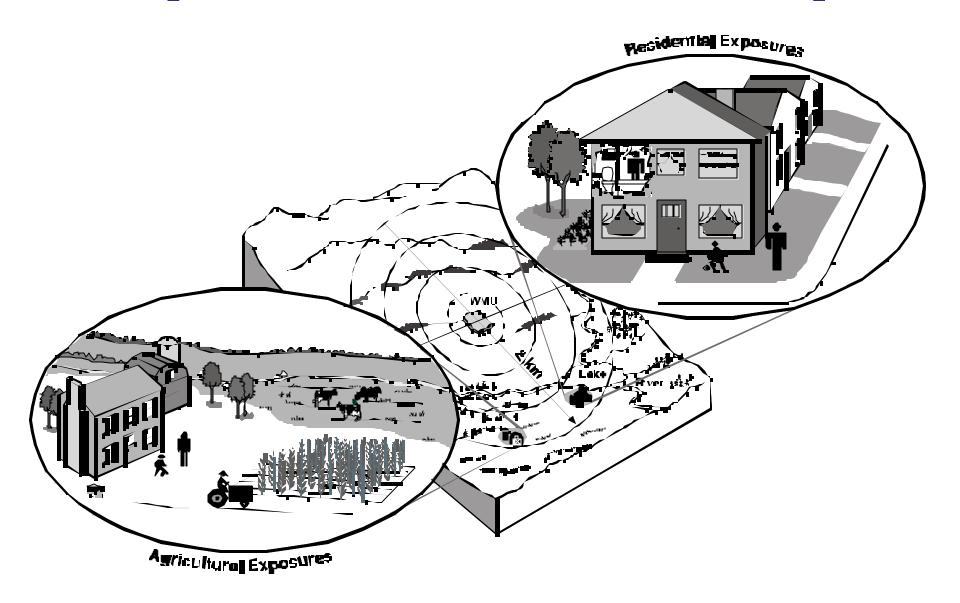
U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

## 3MRA Version 1.0

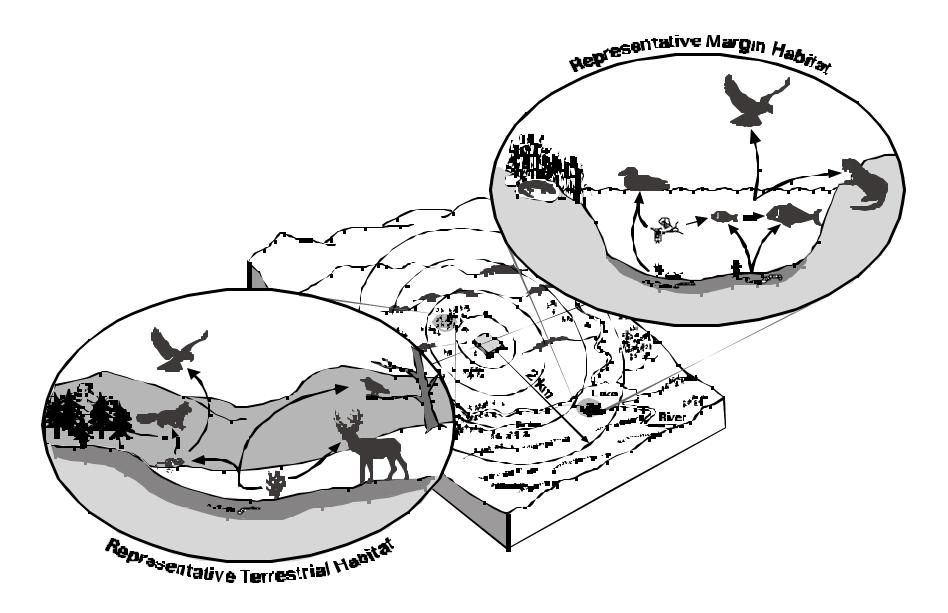
Multimedia, Multipathway, Multireceptor, Risk Assessment

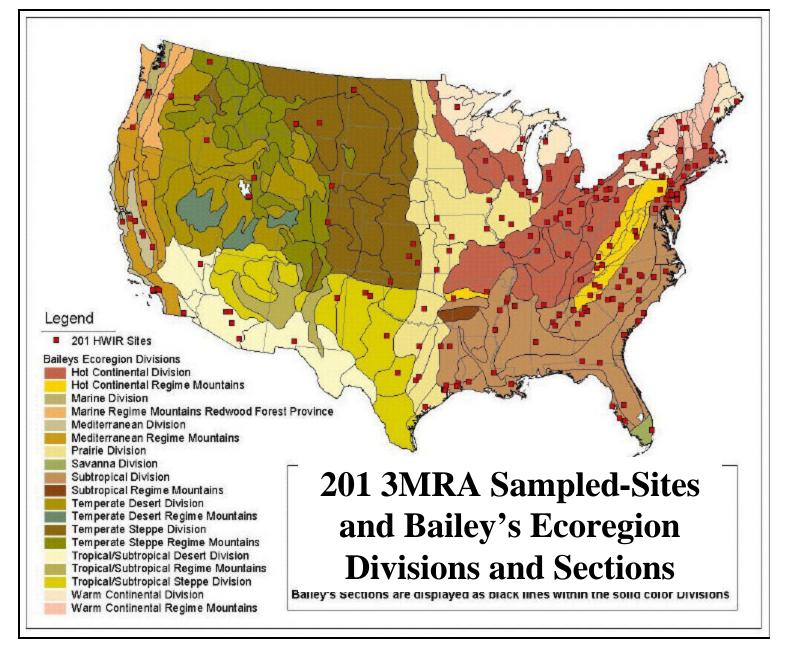
...A screening-level, site-based modeling approach for national-scale assessment of land-based hazardous waste disposal.

#### **Conceptual Framework For Human Receptors**



#### **Conceptual Framework For Ecological Receptors**





#### Overview of 3MRA National Study

#### **Source Types (WMUs)**

- •Surface Impoundment
- Aerated Tank
- •Landfill
- •Waste Pile
- •Land Application Unit

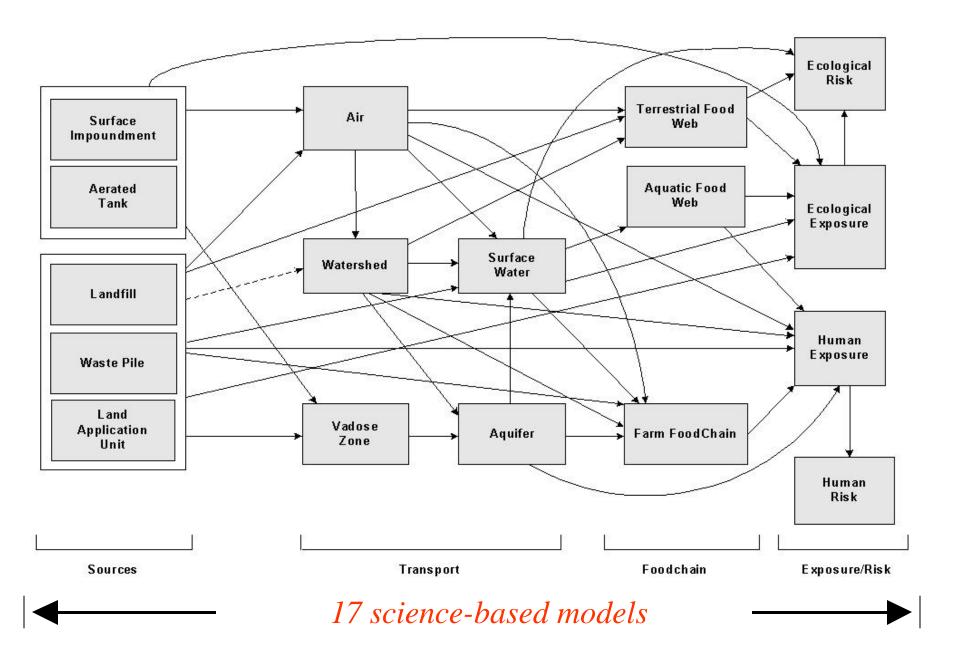
Problem Statement
Conceptual Model
Modeling System
Input Data
Sampling-Based Simulation

#### "Exit Level" Post-Processing

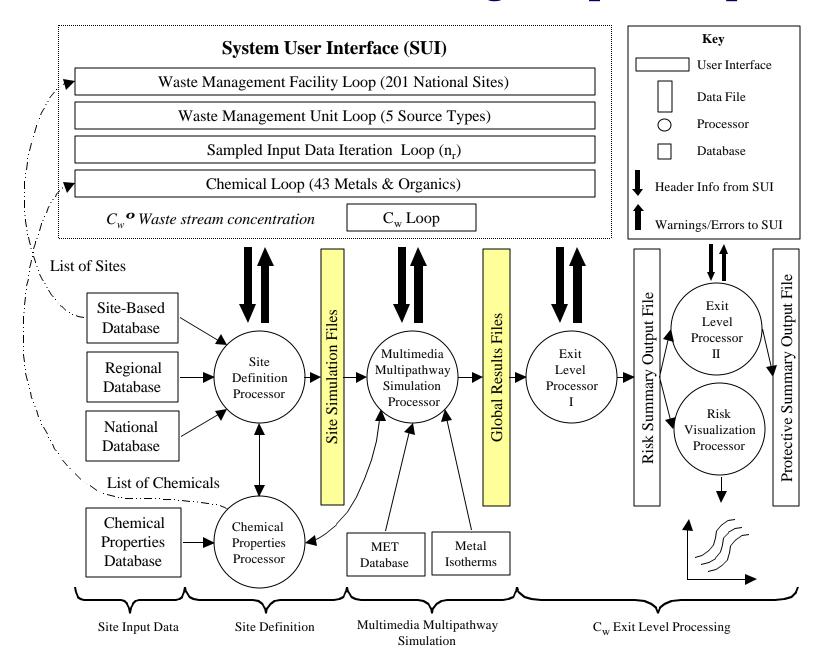
**Output Data** 

- •Variability, Uncertainty and Sensitivity
  - Cancer (risk probability)
  - Noncancer (hazard quotient)
  - •Population weighted risk distribution
    - •Multiple protection measures

#### 3MRA Science Modules and Connectivity



#### **3MRA Stand-Alone Design: Input-Output**



### 3MRA Model Input Variables (966)

SystemGroup	3MRA10	-	
DictCode	ssf	•	
Number of Inputs	55		: ::
ModGroup	▼ Dictionary Description	*	Total
af	aquatic foodweb		21
aq	saturated zone		14
ar	air		23
at	AT		24
ср	chemical properties		142
ee	ecoexposure		14
er	ecorisk	-	4
ff	farm foodchain		45
hd	header		63
he	human exposure		105
hr	human risk		7
lau	LAU		51
1f	LF		44
si	SI		25
s1	site layout		208
sw	surface water		24
tf	terrestrial foodweb		31
vz	vadose zone		3
wp	WP		50
ws	watershed		23
Grand Total			921

Broken down by 20 iconic input dictionaries

Each input - up to 3 dimensions...

### 3MRA Model Output Variables (372)

SystemGroup	3MRA10 🔻	
DictCode	grf	
Number of Inputs		
ModGroup	▼ Dictionary Description ▼	Total
af	aquatic foodweb	18
aq	saturated zone	11
ar	air	17
ee	ecoexposure	3
er	ecorisk	27
ff	farm foodchain	39
he	human exposure	78
hr	human risk	39
s1	site layout	4
sr	source	38
sw	surface water	19
tf	terrestrial foodweb	60
VZ	vadose zone	5
Ws	watershed	14
Grand Total		372

Each output up to 5 dimensions...

These are further post-processed into more useful risk endpoints....





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

# 3MRA National Assessment Strategy

... screening-level, site-based modeling approach for national-scale assessment of land-based hazardous waste disposal.

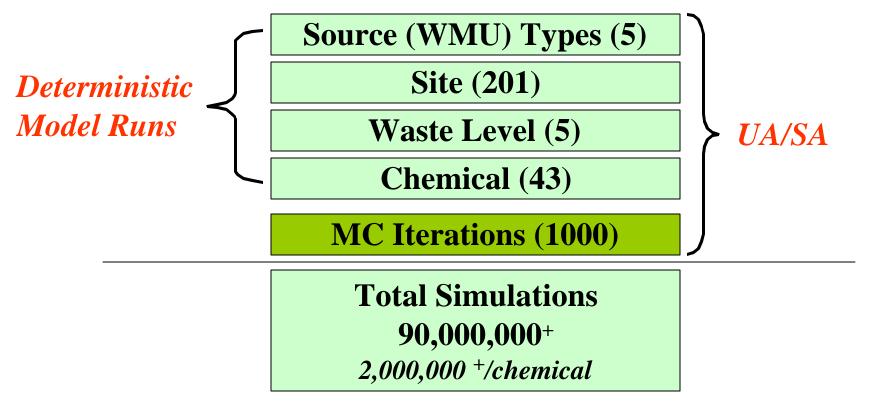
#### **National-Scale Problem Statement**

At what waste stream concentration  $(C_w)$  will wastes, when placed in a non-hazardous waste management unit over the unit's life, result in:

- 1. (Human) Greater than A% of the people living within B distance of the facility with a risk/hazard of C or less, and
- 2. (Ecological) Greater than D% of the habitats within E distance of the facility with an ecological hazard less than F,
- 3. (National) At G\% of facilities nationwide,
- **4.** (Uncertainty) With confidence H% accounting for subjective input uncertainty, and confidence I% accounting for output sampling error.

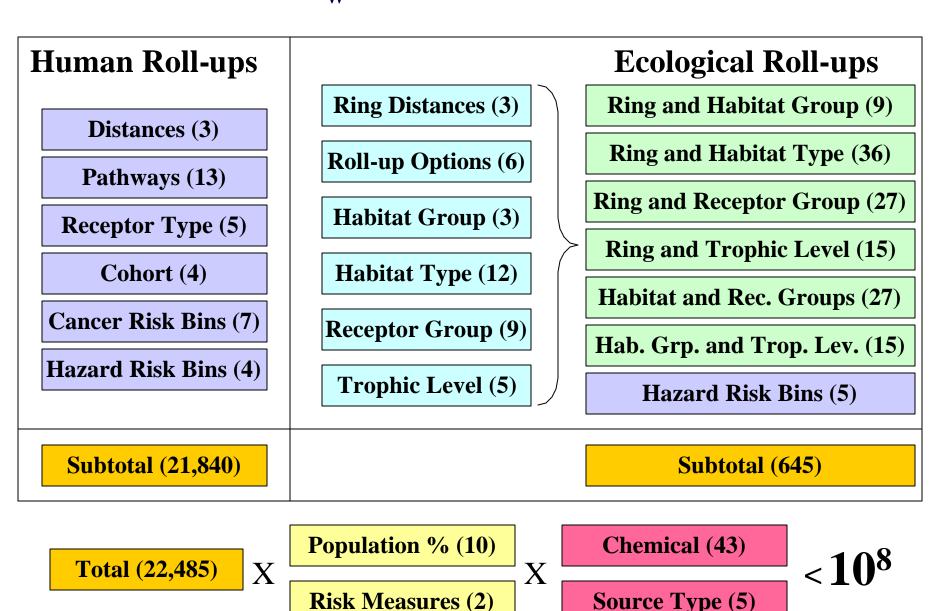
#### Monte Carlo Simulations Needed for 3MRA National Assessment

Individual 3MRA Modeling System Simulations Needed

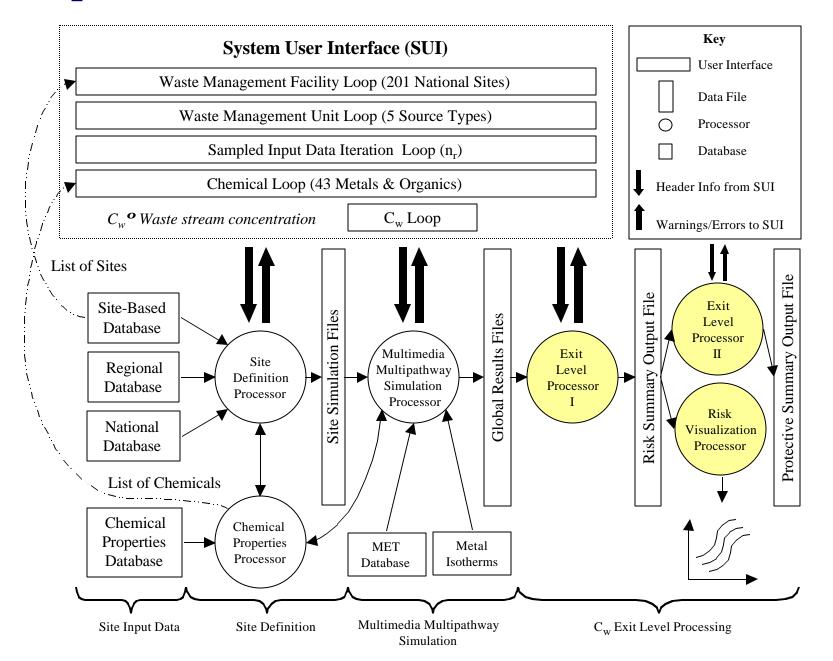


One national iteration for 1 chemical, 1 WMU = 2095 model runs (i.e., for exit levels:  $5 C_w * 419$  site-WMUs = one output sample)

#### Wastestream (C<sub>w</sub>) Exit Levels Possible in 3MRA



#### Importance of the "Exit Level" Processors.....



#### **Problem Statement Revisited**

In determining a single exit level waste stream concentration ( $C_{we vit}$ ):

#### **Output Profile Scenario Description**

a. Chemical

b. Source Type (i.e., WMU)

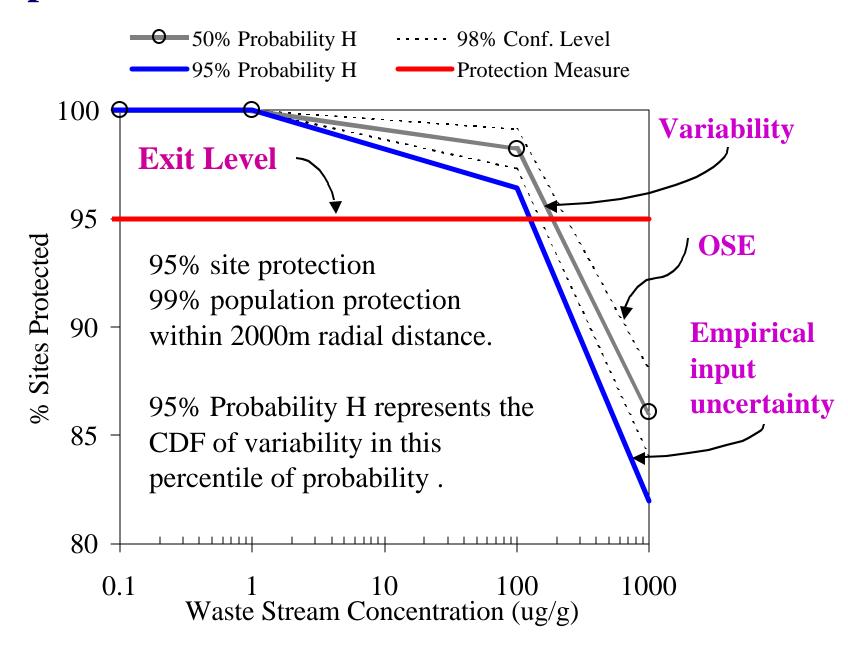
c. 9-tuple Risk Profile:

A% human pop. D% ecol. pop. G% sites

B distance E ring distance H% uncertainty

C risk level F risk level I% precision

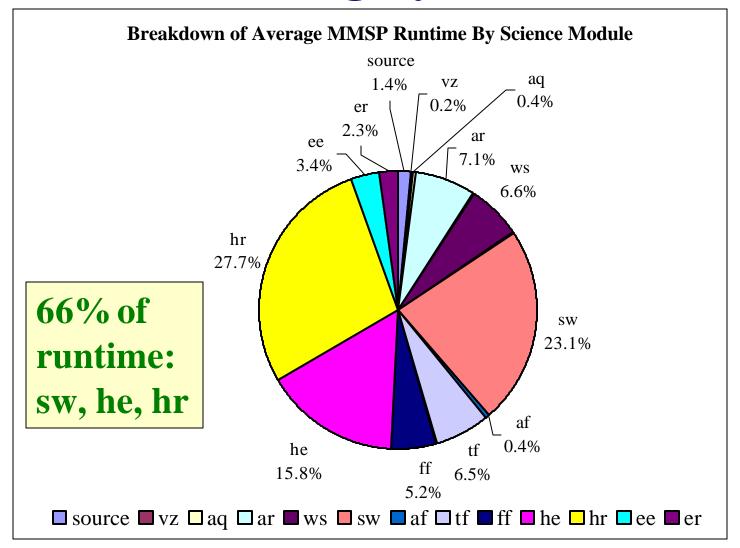
#### **Example 3MRA Exit Level Calculation for Benzene**



#### Summary of Dimensionality for National-Scale Problem Statement

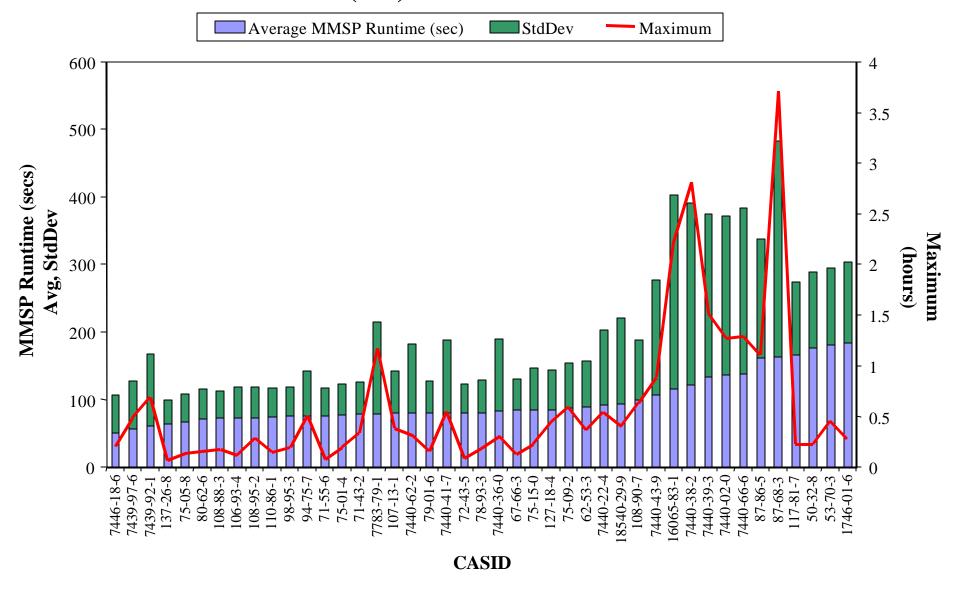
- # Module/Processor Inputs = 966 per model run ( $\leq 3D$ )
- # Module/Processor Outputs = 372 per model run ( $\leq 5D$ )
- Model runs needed for probabilistic national risk assessment per chemical (i.e.,  $UA_p$ ): ~ 2,000,000 (+)
- Post-processed exit levels  $(C_w)$  possible for a 3MRA national assessment:  $\sim 10^8$  accounting for:
  - Multiple decision variables in risk context
  - Population and subpopulation analysis
  - 43+ chemicals, 5 waste management unit types

## 3MRA Modeling System Runtimes



Average individual site scenario run ~ 2 min.

#### 3MRA (1.0) Total MMSP Runtimes



#### 3MRA Run Time for National-Scale Problem Statement

- 1. All chemicals: average model run time ~2 minutes
- 2. On SuperMUSE:  $\sim$  2 days per 100 national realizations of model runs needed for probabilistic national risk assessment per chemical (i.e., UA<sub>p</sub>)
- 3. Single PC: ~ 10 months per chemical per 100 realizations
- 4. Single PC: ~ 8 years per chemical per 1000 realizations
- 5. Single PC: ~ 344 years, 43 chemicals, 1000 realizations





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

# SuperMUSE: Supercomputer for Model Uncertainty and

... 180-client nodes,

215 GHz, PC-Based - Windows and Linux both supported

**Sensitivity Evaluation** 

## Program Server Master Console Model Tasker(s) CPU Allocator **KVM Switches** Network Switches Uses stand-alone version of model software Results Database(s) Server) e.g. mySQL Data Server; Data Analysis Software Hardware

#### **SuperMUSE**

**Supercomputer for Model Uncertainty** & Sensitivity Evaluation

#### **Major Components**:

- Front-end program server,
- Back-end data server,
- Currently 180 client PCs
- 16-port Raritan KVM switches,
- 24-port Linksys (10/100) switches
- Master CISCO 3550-24/2 switch.
- Network protocol TCP/IP.
- GigE channel (1000 megabits/sec) data flow to and from servers.

Windows & Linux OS supported

## SuperMUSE Parallel Computing Cluster at ORD/NERL/ERD, Athens, Georgia

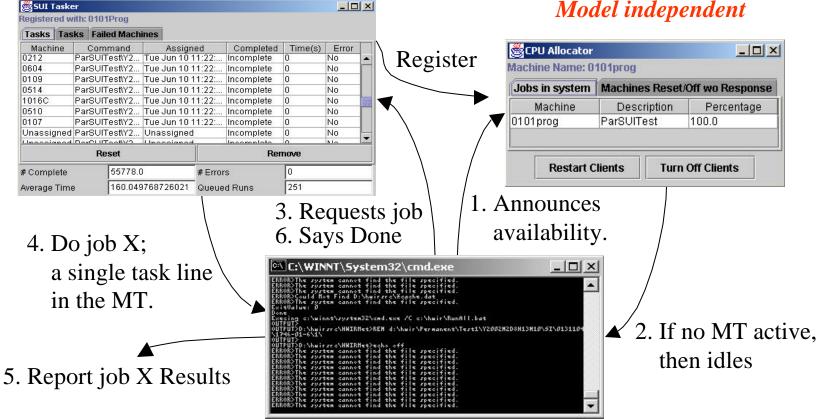


#### **Model Tasker (MT)**



#### **CPU Allocator**





#### **Tasker Client**

Model independent

Executes OS-based commands delivered by MT

#### MiniMUSE Parallel Computing Cluster



#### Beneficial Impacts of PC-Based SuperMUSEing

- ✓ SuperMUSE is scalable from 2 to 1000<sup>+</sup> PCs.
- ✓ Supports Windows or Linux based modeling systems.
- ✓ Solves "embarrassingly parallel" computing problems.
- $\checkmark$  A local solution  $\Rightarrow$  empowers model developers and users.
- ✓ Simple, inexpensive, can be built/operated by PC novices.
- ✓ Ideal for debugging models and performing UA/SA.
- ✓ For an average model runtime of 2 minutes, ERD's SuperMUSE can run over 3 million simulations/month.





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

## 3MRA Version 1.x Tools to Support UA/SA

...includes both model dependent and model independent tools

#### **Supercomputing Software System Needs**

# Facilitating Distribution of Workloads Among PCs

- •CPU Allocator ✓
- •Model Tasker ✓
- •Tasker Client ✓
- •Client Monitor 🗸

## Managing files and Data Across PCs

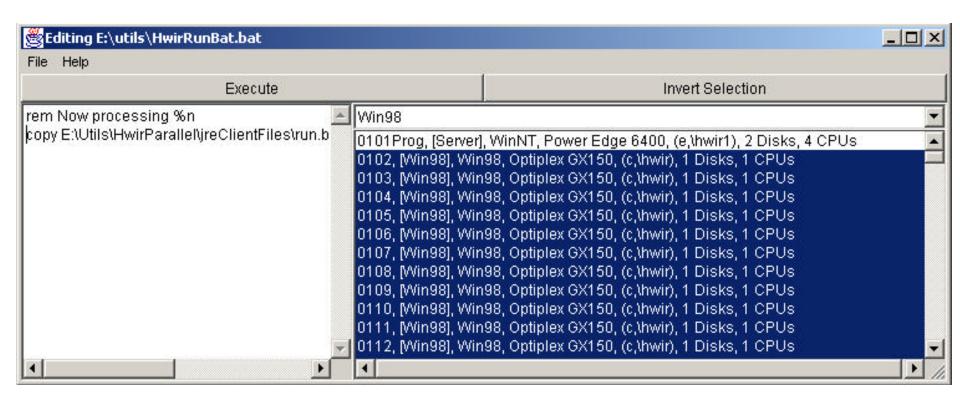
- Update Client 🗸
- Command Tasker 🗸
- Process Messages 🗸
- ELP1 Client Collectors
  - •Aggregated 🗸
  - Disaggregated

## Facilitating 3MRA-Specific Data Analysis

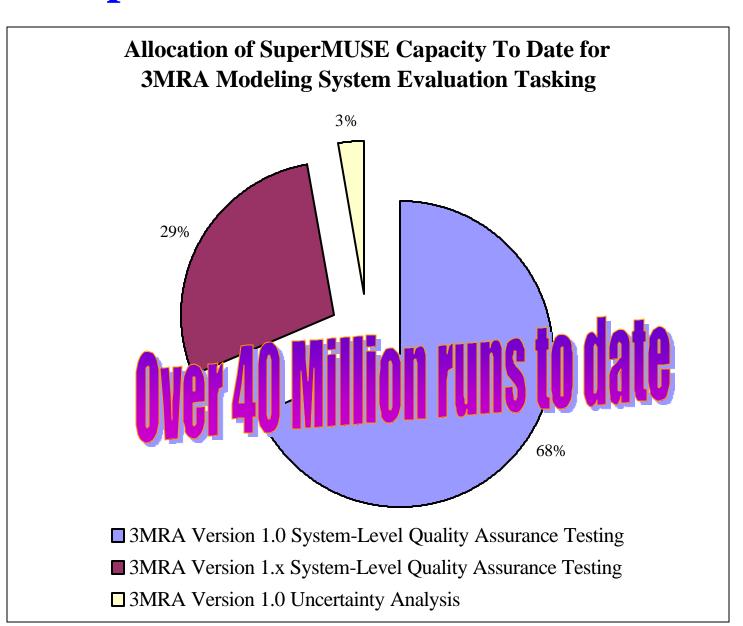
- Site Visualization 🗸
- Site Summary ✓
- Aggr. MySQL ELP1 🗸
- Aggr. MySQL ELP1 🗸

- Aggr. ELP2Vis 🗸
- Automated UA/SA Tools
- Disaggregated ELP1
- Disaggregated ELP2
- Enhanced SUI

# **Example: Server-Side Update Client Tool User Interface**



#### Use of SuperMUSE and 3MRA Version 1.x







U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

# Model Evaluation Approaches

...leading to an overall statement of quality assurance in design for a specific intended purpose

(*Beck et al, 1997*)

## Classes and Types of Uncertainty

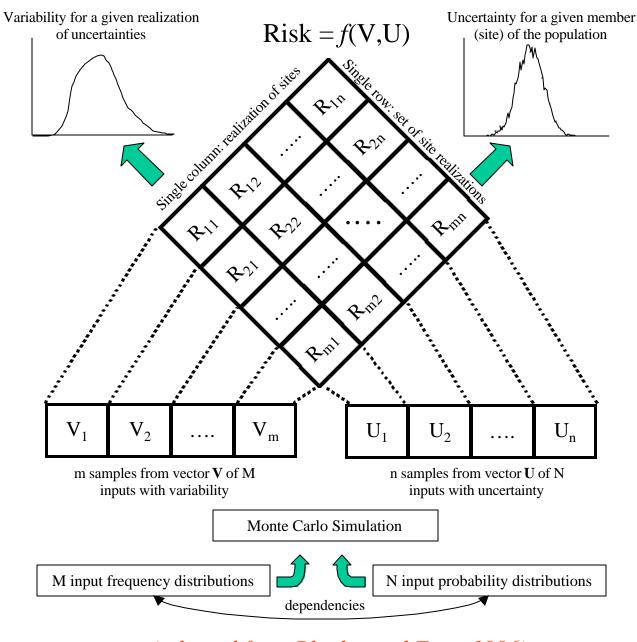
General Classes of Uncertainty									
Variability (V)									
Empirical Uncertainty (U)									
Model Error (ME)									
Types of Empirical Uncertainty									
Random Error (RE)									
Systematic Error (SE)									
Sample Measurement Error (SME; see RE, SE)									
Input Sampling Error (ISE; see RE)									
Output Sampling Error (OSE; see RE)									
Inherent randomness									
Correlation									
Disagreement									

### Performance Uncertainty Analysis (UA<sub>p</sub>)

•••describing potential differences between model predictions and nature.

Uncertainty due to lack of knowledge and data.....

Analysis Approach  $\rightarrow$  given uncertainty in both models and their inputs, quantify/qualify uncertainty in model output(s).

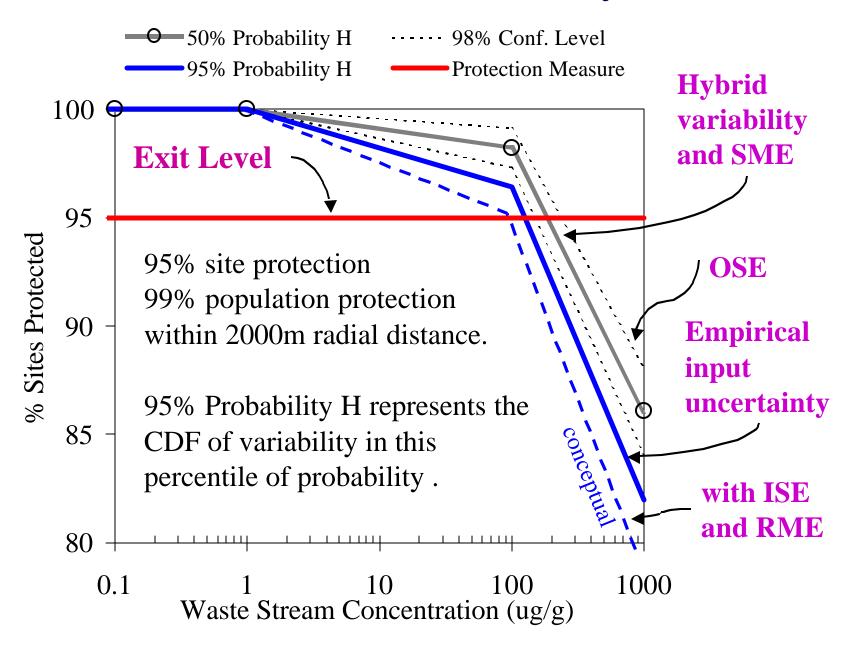


2-Stage
Monte
Carlo....

Separating
uncertainty
and
variability
in output
predictions.

(adapted from Rhodes and Frey, 1996)

#### **3MRA Exit Level Uncertainty**



#### Sensitivity Analysis (SA):

....a study of how the uncertainty in output of an analytical or numerical model can be apportioned to different sources of uncertainty in the model input.

A. Saltelli

#### **Input Space Assessment Techniques**



#### Local

works intensely around a specific set of input values (i.e., the local condition)



#### **Screening**

quick and simplistic, ranks input variables and ignores interactions between variables



#### Global

quantifies scale & shape of the I/O relationship; all input ranges; assesses parameter interaction

#### Components of Model Evaluation.

(a.k.a., Verification, Validation, and Predictive Uncertainty Analysis)

Components of Model Evaluation							
Uncertainty (U)							
Variability (V)							
Total Uncertainty (TU)							
Compositional Uncertainty Analysis (UA <sub>c</sub> )							
Performance Uncertainty Analysis (UA <sub>p</sub> = UA)							
Sensitivity Analysis (SA)							
Calibration (CAL)							
Code Verification (CodVer)							
Model Comparison (ModComp)							
Compositional Validity (CompVal)							
Performance Validity (PerfVal)							
Model Validation (ModVal)							
Peer review							

### Sensitivity-Analysis-Based Performance Validation (SA<sub>p</sub>)

...assessing and describing the behavioral and non-behavioral characteristics of the modeling system.

A specific construction of performance validation as a reflection, by way of sensitivity analysis, of the evaluation of the external definition of the current task back onto the internal composition of the model.

# Model Evaluation: Summary Perspective

Model evaluation is seen as a statement of quality assurance in design (i.e., tool or technology):

....a result of the model synthesis and analysis effort,

.....viewed as the outcome of the overall model validation effort for the specific task defined.





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

## 3MRA Version 1.0 UA/SA Plan

.....for the national-scale assessment problem statement

#### 3MRA UA/SA Plan

- Performance uncertainty analysis (UA<sub>p</sub>)
- Sensitivity analysis (SA)
- Sensitivity-based Performance Validation  $(SA_p)$
- 7 Chemicals:
  - Benzene,
  - **PCE**,
  - 2,3,7,8-TCDD,
  - Benzo(a)pyrene,
  - Arsenic,
  - Nickel, and
  - Divalent Mercury.

## 3MRA UA/SA Plan (UA<sub>p</sub>)

#### **Performance Uncertainty Analysis:**

o Entails propagation of input uncertainty through the modeling system, while also addressing output sampling error (OSE) associated with computational limitations of the sampling-based MCS strategy.

o Uses a pseudo 2<sup>nd</sup>-order analysis to separate variability and empirical input uncertainty, while quantifying OSE.

#### 3MRA UA/SA Plan (SA)

#### **Sensitivity Analysis:**

- o A balanced, tiered formulation of SA is planned for identifying key, important, and redundant model inputs. The basic approach to be undertaken is global input space assessment via sampling-based methods.
- o SA to be conducted for this purpose will enhance both compositional and performance validation aspects for the modeling system.
- Will include global based SA techniques:

Correlation / Regression
Regional Sensitivity Analysis (RSA)
Tree Structured Density Estimation (TSDE)

## 3MRA UA/SA Plan (SA<sub>p</sub>)

#### **SA-based Performance Validation:**

o Basically an assessment of a "prior" validity through the execution of a regional sensitivity analysis (RSA) procedure, realized as an assessment of the model's maximum relevancy in predicting model behavior for various population percentiles.

Will investigate higher order interactions via TSDE.





U.S. EPA Office of Research and Development National Exposure Research Laboratory Ecosystems Research Division

# Example 3MRA Version 1.x Output: 79 National Realizations

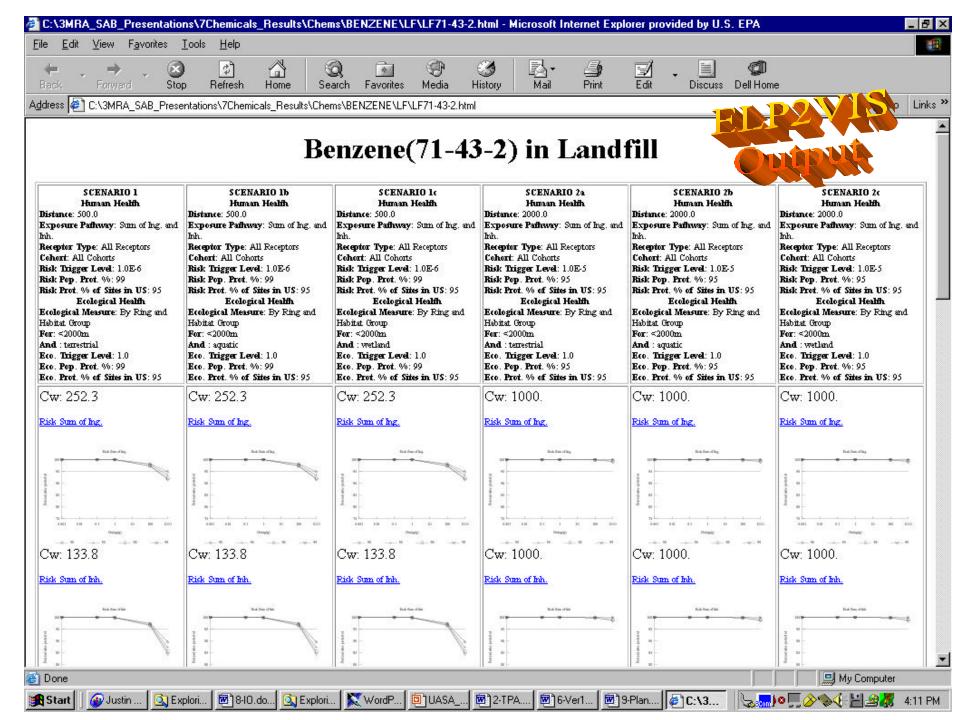
Risk Endpoint Description	Scenario					
Scenario ID	1	2				
% Population Protected	99%	95%				
% Sites Protected	95%	95%				
Protective?	More Less					
Human						
Distance (m)	500	2000				
Cancer Risk	$10^{-6}$	$10^{-5}$				
Hazard Risk	0.1	1				
Pathway	Sum Ing. & Inh.					
Receptor Group	All					
Cohort Group	All					
Ecological						
Ring Distance (m)	2000	2000				
Hazard Risk	1	1				
By Ring and Habitat Group	Terr., Aq., Wetland					
Simulation						
Sources/Chemicals/Cws	5/7/5					
Sites/sources	419					
# National Realizations	79					
# Modeling System Runs	1,158	3,535				

# Example Simulation Experiment

#### Simulated 7 Chemicals & Metals

CASID	Chemical Name						
		Carci	noge	enic	Non- Car	Ecological	
		Inhalation	ı I	ngestion	Inhalation	Ingestion	Zeorogreur
127-18-4	PCE	<b>√</b> (	$\bigoplus$	✓		✓	<b>√</b>
71-43-2	Benzene	✓ (	$\bigcirc$	✓			✓
1746-01-6	2,3,7,8-TCDD	✓ (	$\bigcirc$	✓		✓	✓
50-32-8	Benzo(a)pyrene			✓			✓
7440-38-2	Arsenic	✓		✓		✓	✓
7440-02-0	Nickel	✓				✓	✓
7439-97-6	Dival. Mecury				✓	✓	✓

→ Additive Risk



# **Example Exit Level Results**2 Scenarios - 7 Chemicals

(Example, Preliminary Summary)

Source Type	Suri Impour		Aerated Tank		Land Application Unit		Waste Pile		Landfill	
#Sites Evaluated	137		137		28		61		56	
Total Simulations	378,	805	378	,805	77,420		168,665		154,840	
	50 <sup>th</sup> Percentile (Average) Exit Level (ppm)								-	
Chemical/Scenario	1	2		2		2		2	1	2
PCE	0.36	100*	6.5	100*	3.2	5000*	453	5000*	778	5000*
Benzene	0.13	100*	1.5	100*	2.1	1000*	8.1	1000*	109	1000*
Benzo(a)pyrene	0.001*	0.001*	0.001*	0.001*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*
2,3,7,8-TCDD	2.2E-7	1E-4*	1E-4*	1E-4*	1.4E-6	1.6E-5	2.5E-4	0.001	0.001	0.004
Arsenic	0.095	2.1	50*	50*	0.004	0.59	3.0	18	1.8	298
Nickel	11	1000*	1000*	1000*	6.7	127	352	10000*	446	10000*
Divalent Mercury	0.1*	0.1*	0.1*	0.1*	0.49	2.2	10	10*	10*	10*

\* Shows max Cw analyzed in experiment

**Scenario** 1 = more protective

**Scenario 2 = less protective** 

# Dominant Human Pathways, Human Receptors, and Ecological Habitats

(Example, Preliminary Summary)

Chemical	Dominant Risk/Route	200000000000000000000000000000000000000	face ndment	Aerated Tank		Land Application Unit		Waste Pile		Landfill	
Scenario		1	2	1	2	1	2	1	2	1	2
PCE	HCancer	Cr, Wa	Crop	Стор	Стор	Water	Crop	Стор	Crop	Стор	Стор
	Ing.	Fa,G,Fi	Fa,G,Fi	G,Fi	G,Fa,Fi	G,R,Fi	G,Fi	G,Fi,R	G,Fi,Fa	Fa,G,Fi	Fa,G,Fi
Benzene	HCancer	Air,Cr	Air, Cr	Air,Cr	Air,Cr	Air, Sh.	Air	Show.	Air	Show.	Air
1900.000	Inh., Ing	G,Fi,R	G,Fi,R	G,Fi,R	G,Fi,R	G,Fi,R	G,Fi,R	Fi,R,G	Fi,R,G	G,Fi,R	All
Benzo(a)pyrene	Low Risk	Cr, Beef		7 34		Стор		Crop	1 2	Bf,Cr	5 SHS
	3773	( <del>10</del> 8		0.55	750	1.00			3553	4 9 <del>.5</del> 5	S-5
2,3,7,8-TCDD	HCancer, Eco	Air	Air	Air	3	Air	Eco	Fi,Cr,Air	Eco	Eco	Eco
	Inh.	Bf, Oth.	, <del>2</del>	Bf,Oth.		All	Ter,Aq	Fi,G,R	Wet.	Ter,Aq	Ter,Aq
A SERVICE AND SERV											
Arsenic	HCancer, Eco	Water	Eco	1044	94.0	Wa, Cr	Eco	Wa,Cr	Eco	Water	Eco
	Ing.	Fi	Ter.	· ·		G,Fi,R	Aq.,Ter	G,Fi,R	Wet.	All	Aq.
Nickel	HBazard, Eco	Water	99	S		Water	Eco	Water	Eco	Water	Eco
	Ing.	Fisher				R,G,Fi	Ter,Aq	R,G,Fi	Wet.	Fa,Fi	Aq,Ter
Hg <sup>-7</sup>	HBazard	122	<u> </u>		220	Fish	Fish	Fish	Fish	Eco	Eco
	Ing.					Fisher	Fisher	Fisher	Fisher	Ter,Aq	Ter,Aq